Best Practices and Controls for Mitigating Insider Threats

George Silowash
Team Member, Technical Solutions

- Digital Forensic Investigations & Incident Response
- Information Assurance Risk Management
- Open Source Solutions

Alex Nicoll
Team Lead, Technical Solutions

- Information Assurance
- Operating System Design
- High Assurance Systems (MLS)

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
08 AUG 2013		**00-00-2013 to 00-00-2013**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Best Practices and Controls for Mitigating Insider Threats	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Carnegie Mellon University,Software Engineering Institute,Pittsburgh,PA,15213**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

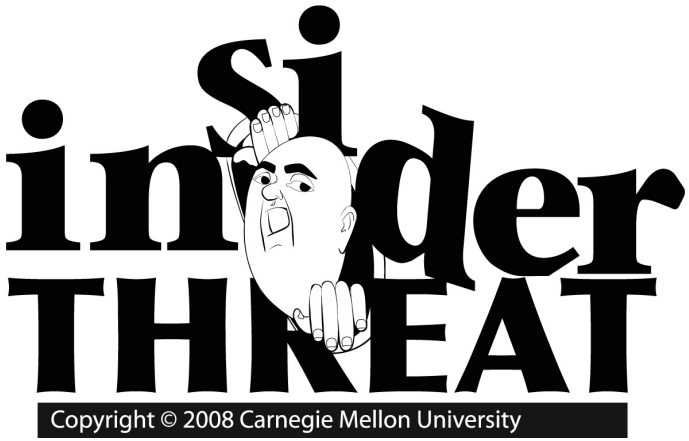
14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**53**	

Agenda

- Introduction
- Common Sense Guide to Mitigating Insider Threats, 4th Edition
  - 19 Best Practices
- Technical Demonstration(s)



CERT Insider Threat Center—Mission

Assist organizations in identifying indications and warnings of insider threat by

- performing vulnerability assessments
- assisting in the design and implementation of policies, practices, and technical solutions

based on our ongoing research of hundreds of actual cases of insider IT sabotage, theft of intellectual property, fraud, and espionage

Definition of Insider Threat

The CERT Program's definition of a malicious insider is a current or former employee, contractor, or business partner who meets the following criteria:

- has or had authorized access to an organization's network, system, or data
- has intentionally exceeded or intentionally used that access in a manner that negatively affected the confidentiality, integrity, or availability of the organization's information or information systems

Methods

- Research
- Empirical Evidence
- Contarol Hypothesis
- Control Implementation and Testing
- Control Pilot
- Revisions
- Release

Common Sense Guide to Mitigating Insider Threats, 4th Edition

Who does the CSG apply to?

- Information Technology / IT Security
- Physical Security
- Software Engineering
- Data Owners
- Legal
- Human Resources
- …..everyone across the organization

New Features

- Mappings to other best practices / standards
  - NIST 800-53
  - ISO 27002
  - CERT RMM
- Quick wins & High Impact Solutions
- Quick reference guide

Practices you are familiar with

Consider threats from insiders and business partners in enterprise-wide risk assessments.	Consider insider threats in the software development life cycle.
Clearly document and consistently enforce policies and controls.	Use extra caution with system administrators and technical or privileged users.
Institute periodic security awareness training for all employees.	Implement system change controls.
Monitor and respond to suspicious or disruptive behavior, beginning with the hiring process.	Log, monitor, and audit employee online actions.
Anticipate and manage negative workplace issues.	Use layered defense against remote attacks.
Track and secure the physical environment.	Deactivate computer access following termination.
Implement strict password and account management policies and practices.	Implement secure backup and recovery processes.
Enforce separation of duties and least privilege.	Develop an insider incident response plan.

New Best Practices

- Practice 9: Define explicit security agreements for any cloud services, especially access restrictions and monitoring capabilities.
- Practice 16: Develop a formalized insider threat program.
- Practice 17: Establish a baseline of normal network device behavior.
- Practice 18: Be especially vigilant of emerging social media trends.
- Practice 19: Close the doors to unauthorized data exfiltration.

Practice 9

Define explicit security agreements for any cloud services, especially access restrictions and monitoring capabilities.

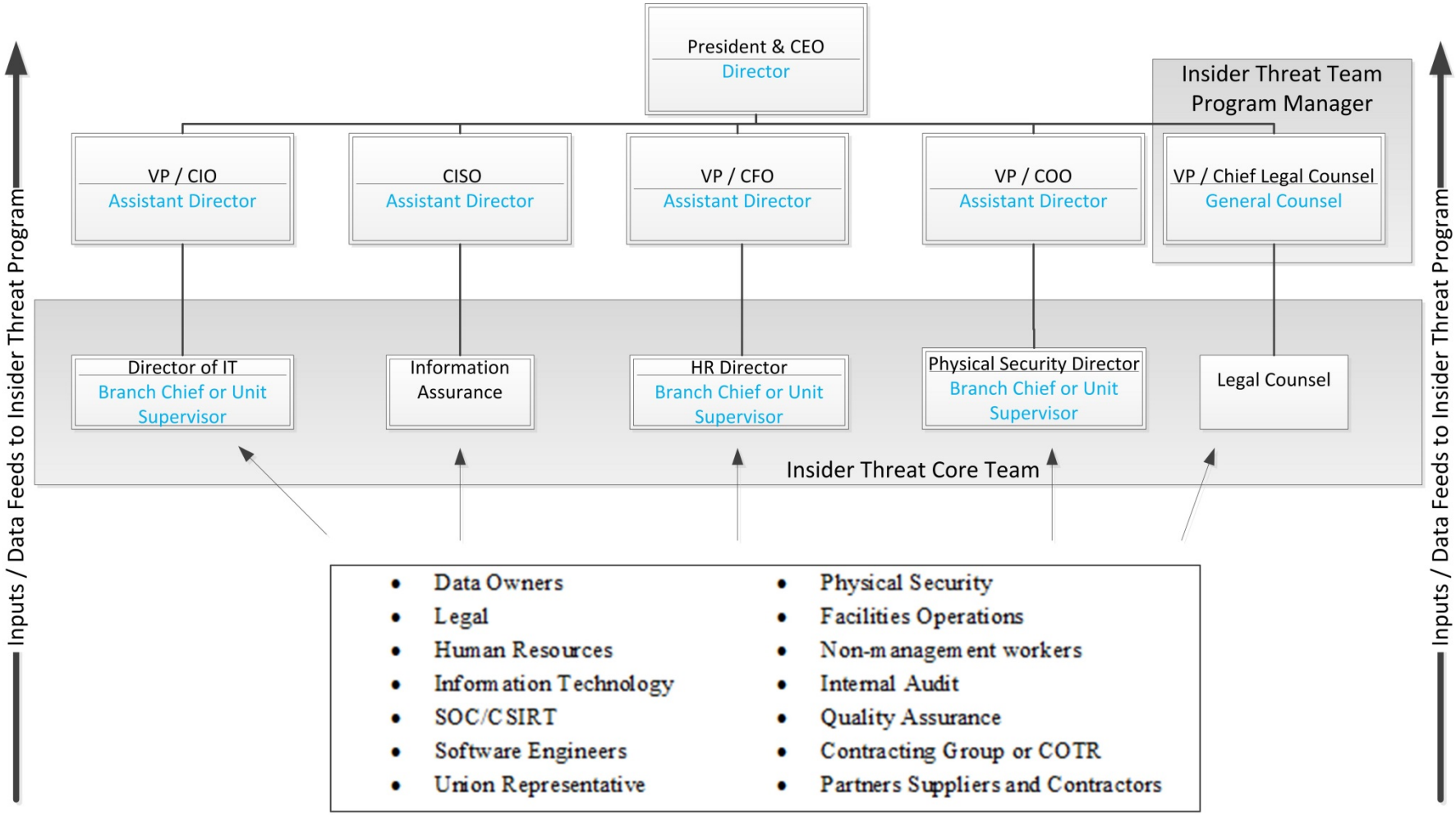
- Conduct a Risk Assessment before entering into any agreement.
- Chose a cloud service provider that meets or exceeds the organization's own levels of security.
- Understand how the cloud provider protect data and other assets.

Practice 16

Develop a formalized insider threat program.

- Work with Legal Counsel.
- Requires involvement from various departments across the organization.
- Share information.

Insider Threat Team



Note: Text below the separator in each box notes the federal government's equivalent position

Practice 17

Establish a baseline of normal network device behavior.

- Know what is normal and abnormal for a given system.
- Excessive traffic, Insufficient traffic
- Store logs for 60 days or longer

Practice 18

Be especially vigilant regarding social media.

- Train users to be aware of what they post
- Small disclosures of information can create bigger problems
- Develop a social media policy

Practice 19

Close the doors to unauthorized data exfiltration.

- Understand how data can leave the organization.
- Control removable media.
- Watch for “old school” methods: printers, copiers, etc.

Technical Controls:

Preventing Data Exfiltration

The Problem

- Organizations need to use web based services on a daily basis for business needs. However, services that offer the ability to upload attachments present an opportunity for sensitive data to leave the organization.

- Communications that are secured with SSL encryption are difficult to inspect and therefore it is difficult to detect and prevent sensitive data from leaving the organization.

Data Loss Through the Web

Difficult problem

Perfect exfiltration channel

- Encrypted
- Appears “normal”
- Send many files at once
- Possibly essential to operations

What can be done to prevent this?

Options:

1. Implement policies regarding how sensitive information is disseminated
2. Full packet capture of all Internet traffic for further analysis
3. White listing
4. Block all webmail services
5. Allow all webmail services and cross your fingers
6. Or…

CERT's Solution

- Allow proxied Internet access to any website
- Inspect encrypted communication sessions for sensitive documents
- Block sensitive attachments from being uploaded to the Internet

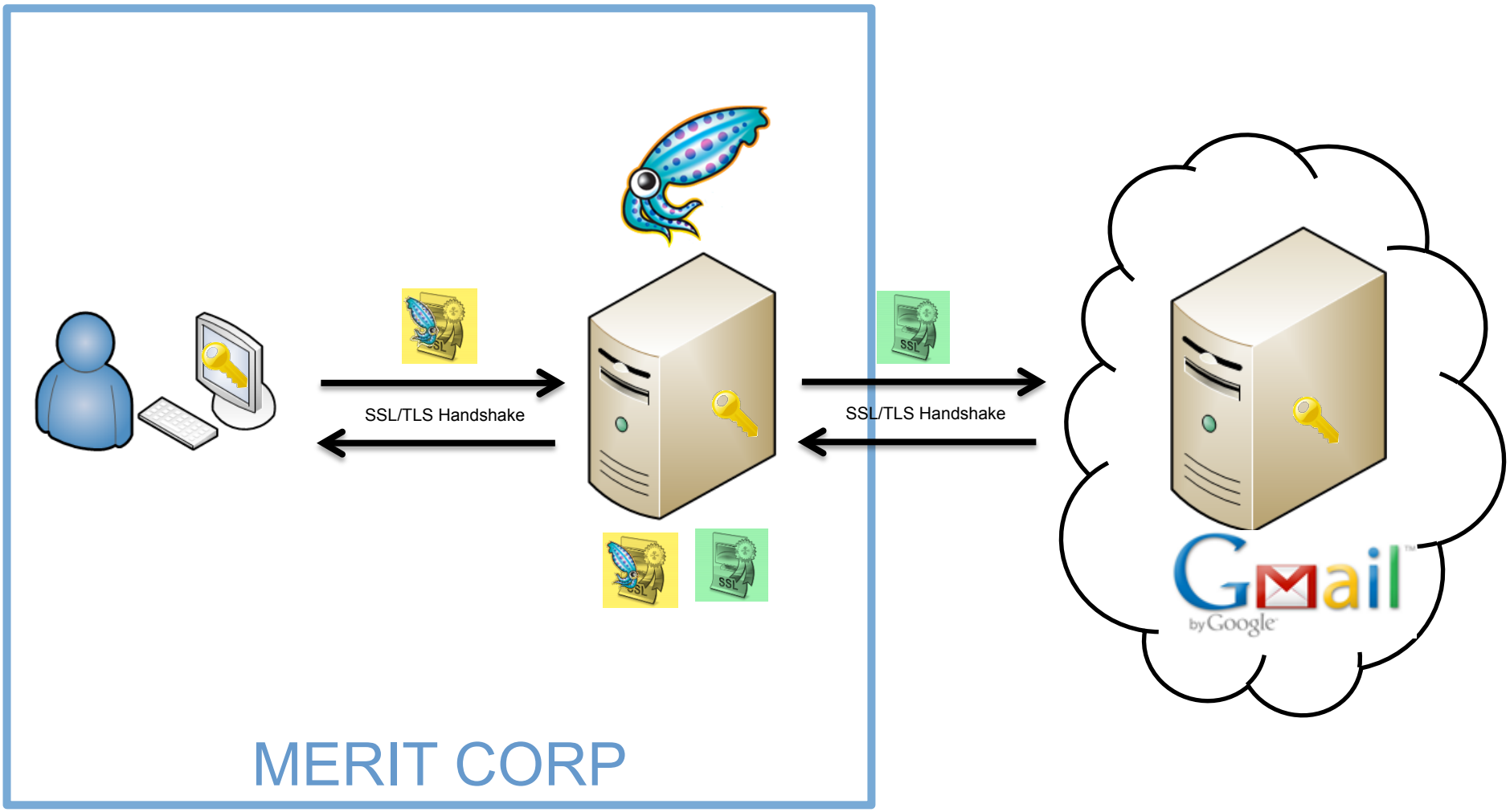
Blocking Documents

Documents can be stopped based on three methods:

1. Block all attachments
2. Keywords
3. Tags

The Proxy Server

Man-in-the-Middle (MITM) Proxy



The Proxy Server Main Components

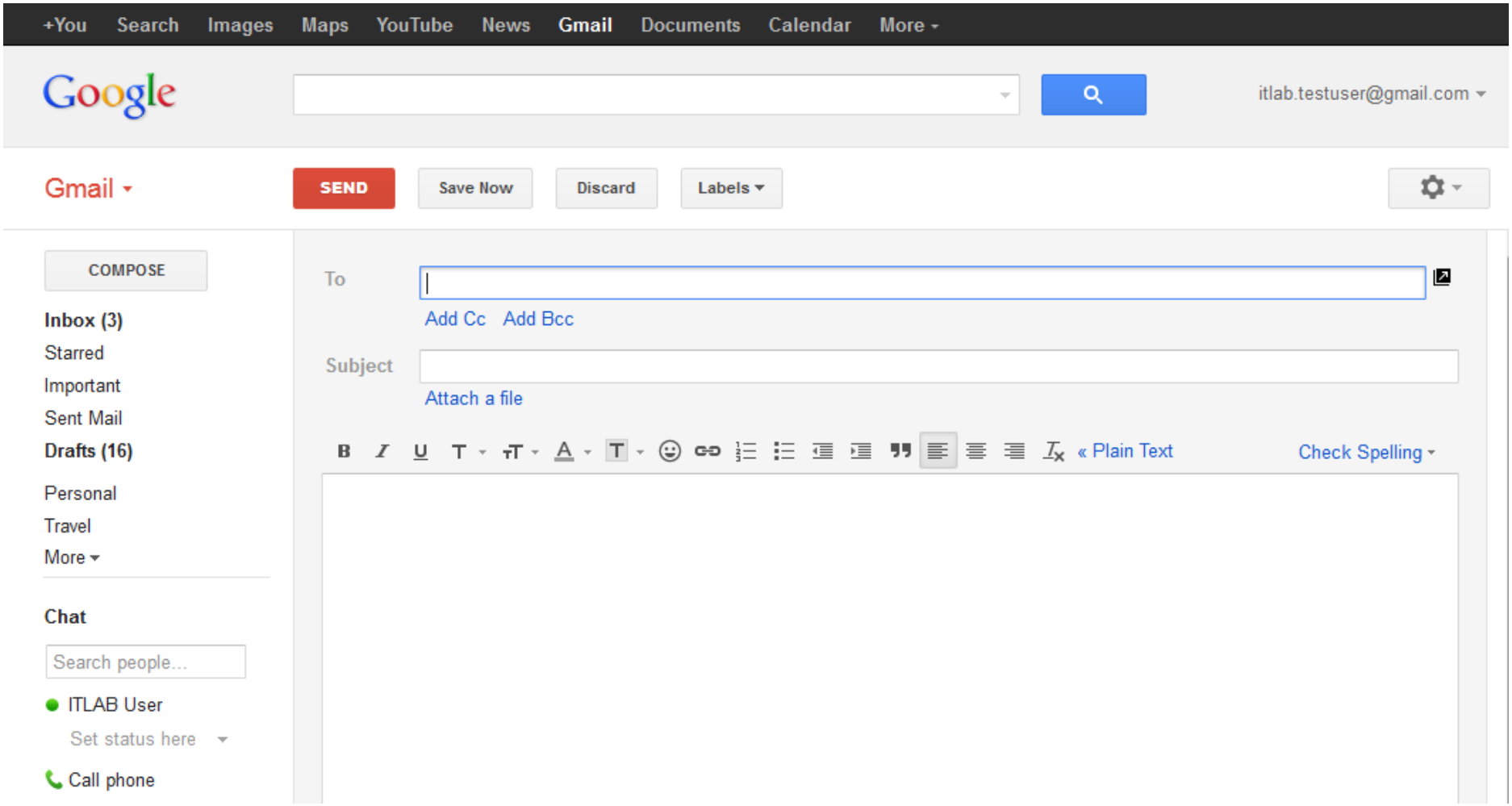
- Ubuntu Linux Version 10.04 LTS
- Squid Version 3.1.19
- C-ICAP
- Clam Antivirus (ClamAV)

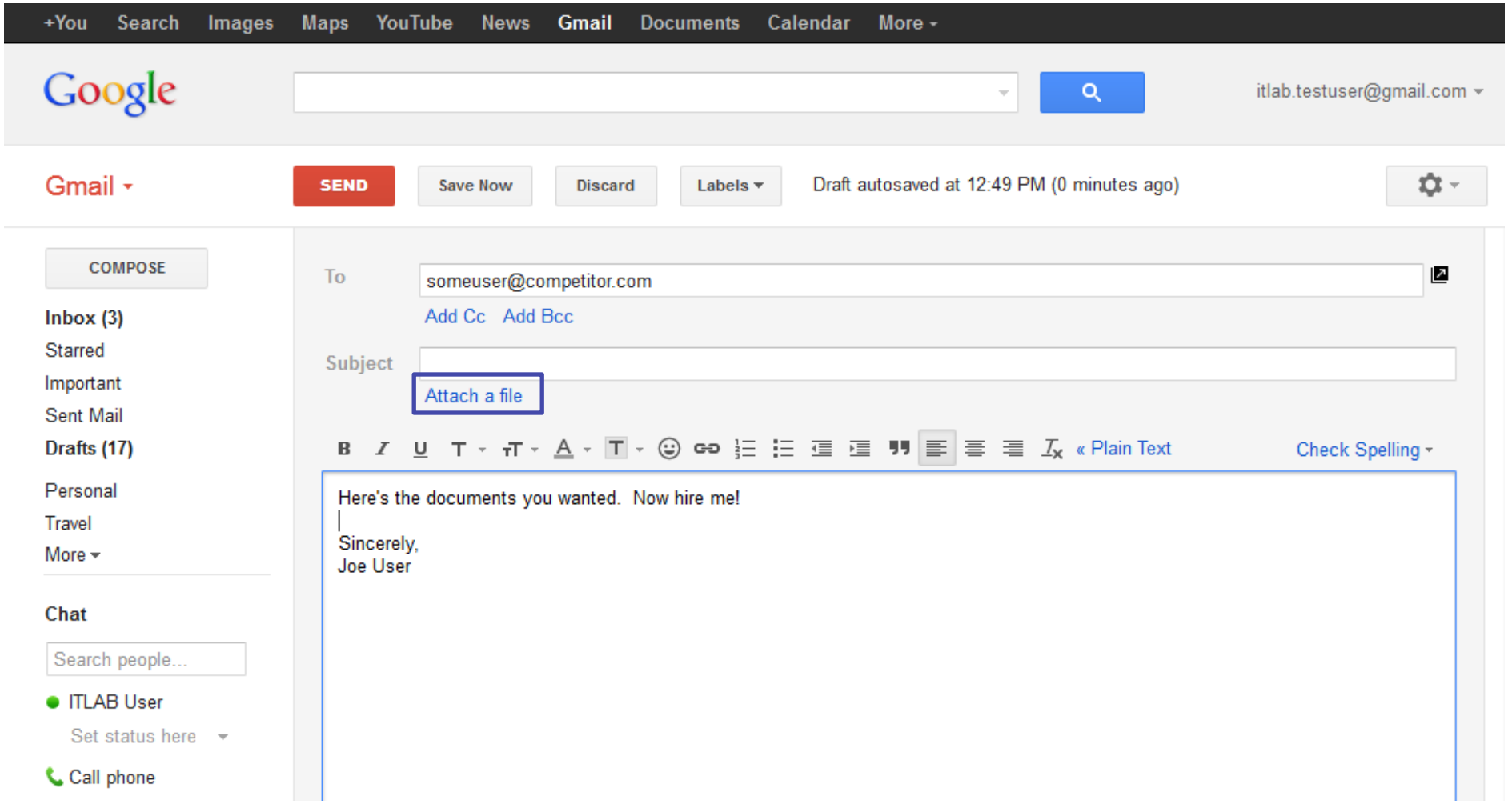
Client Configuration

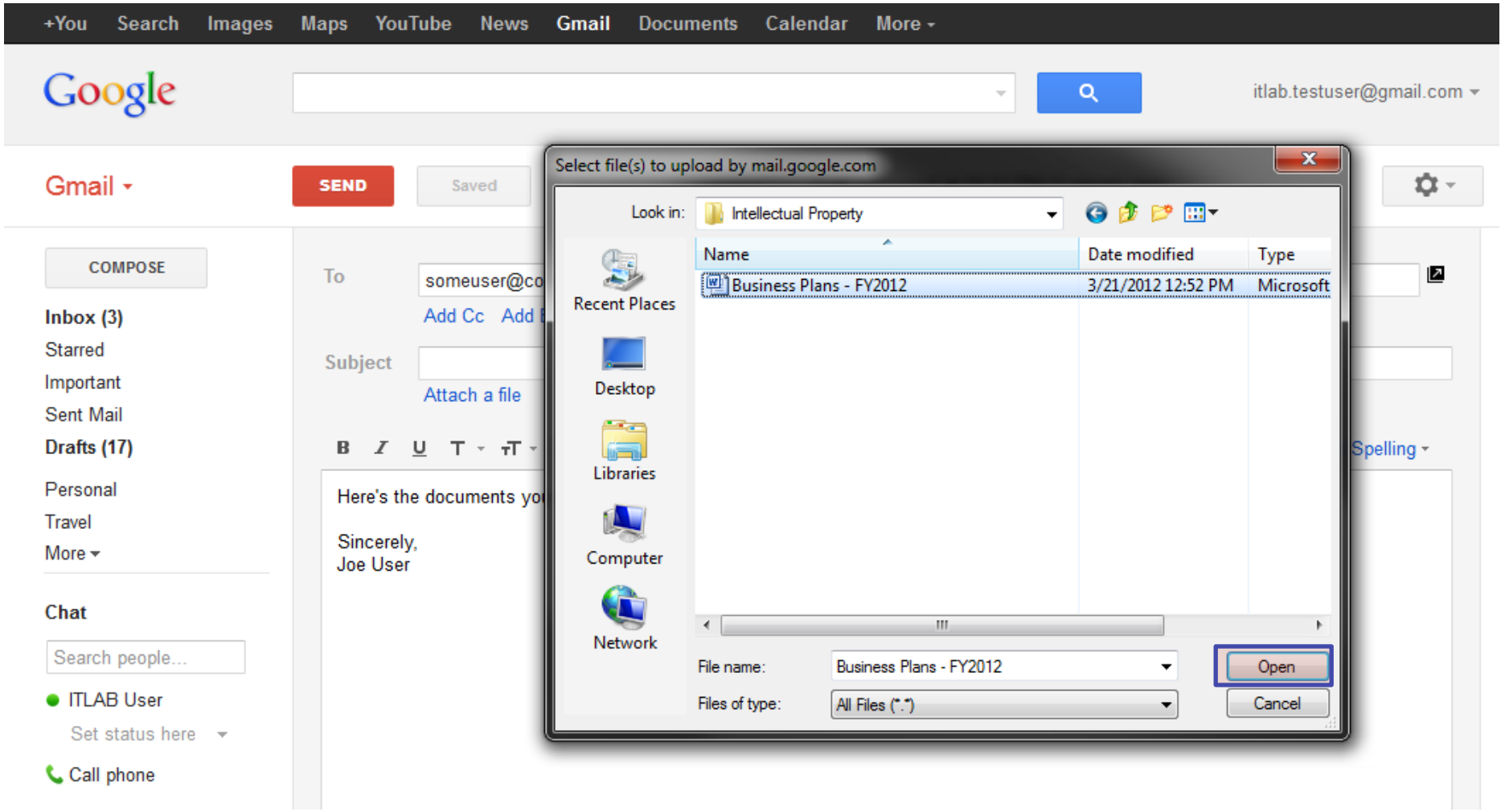
- The Organization's Certificate needs installed in the Trusted Root Certificate Store on each client
- Internet Explorer needs to be configured to use the proxy on port 3128 for HTTP/S traffic

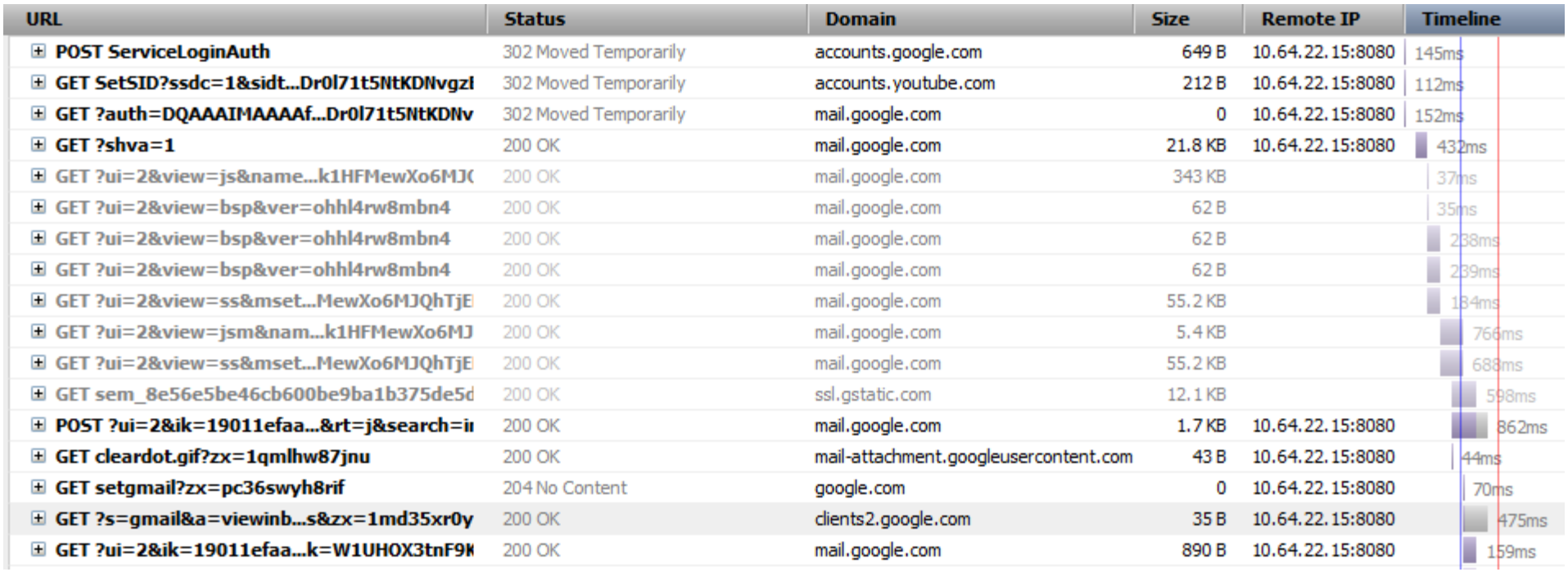
Both of these settings can be configured using

Group Policy

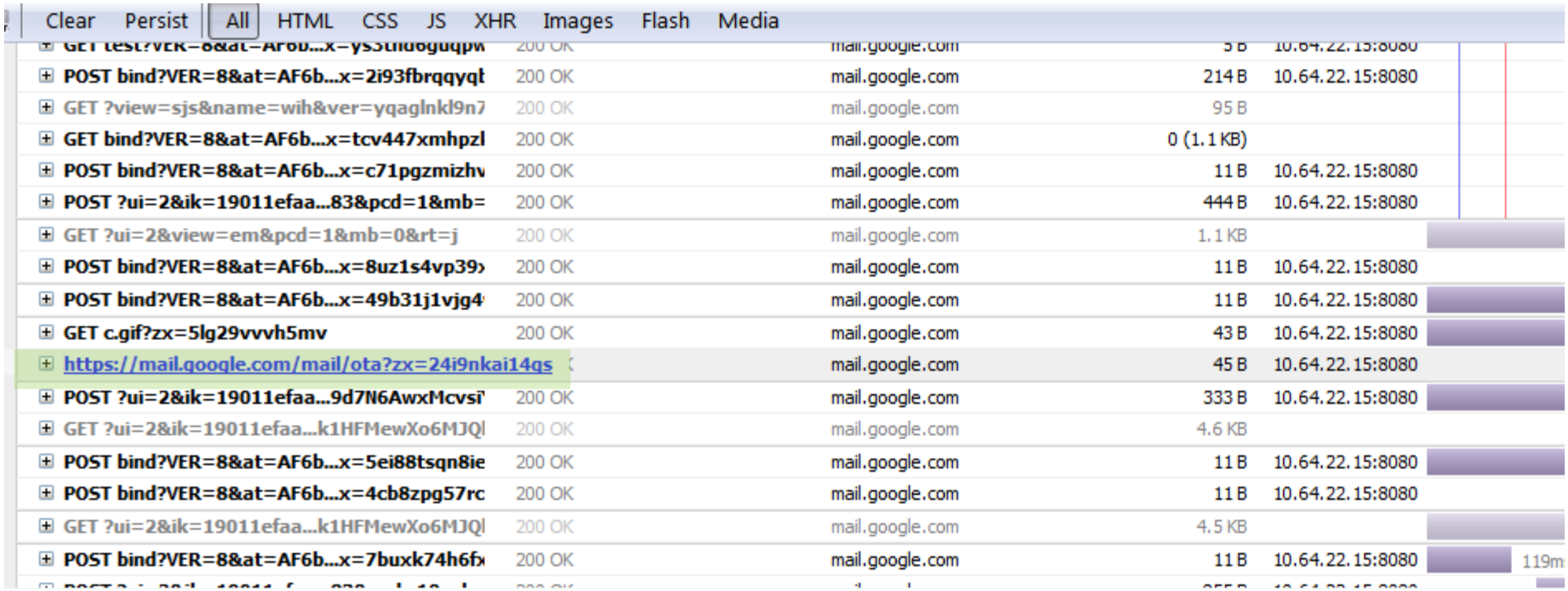




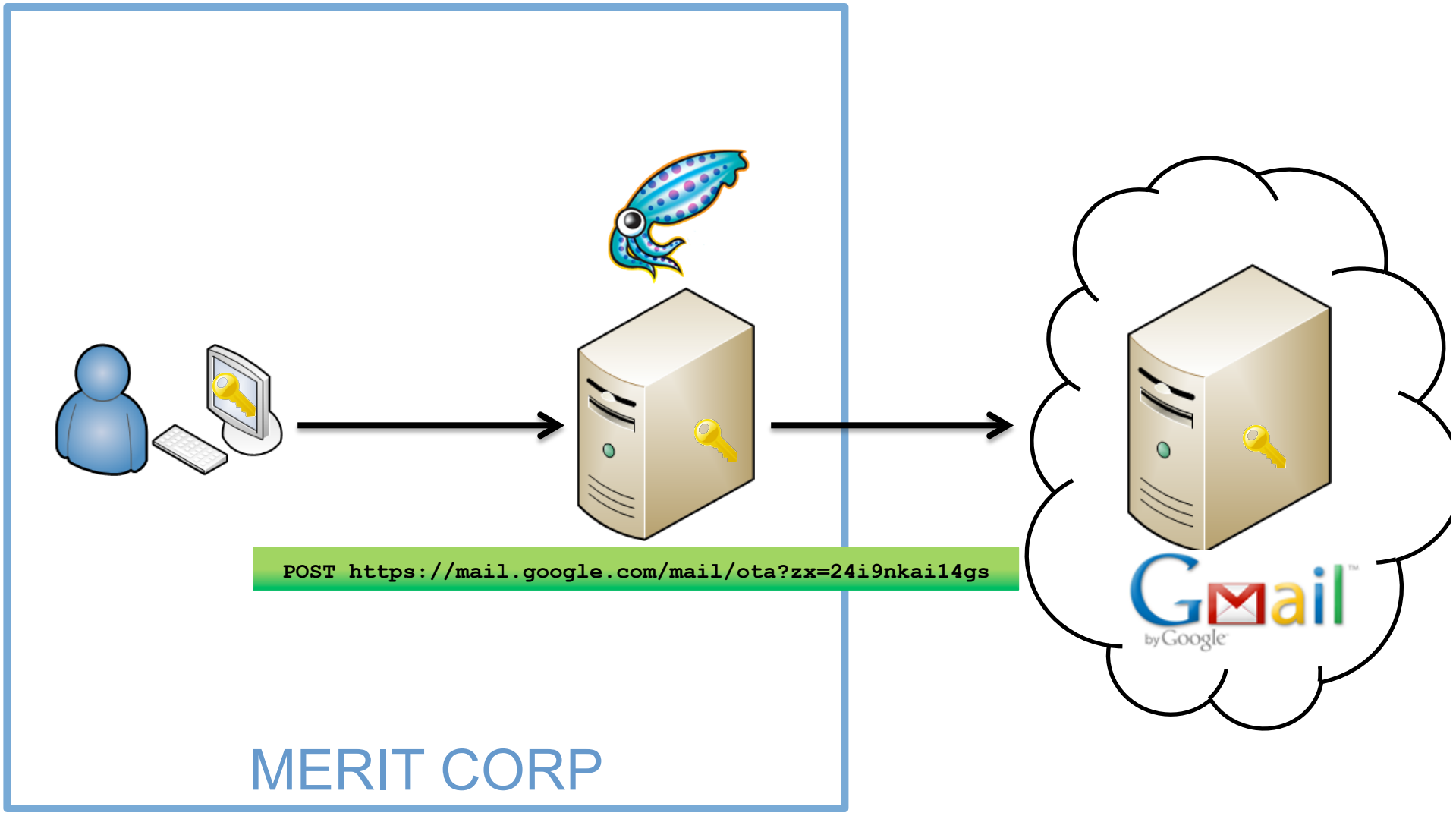




URL	Status	Domain	Size	Remote IP	Timeline
POST ServiceLoginAuth	302 Moved Temporarily	accounts.google.com	649 B	10.64.22.15:8080	145ms
GET SetSID?ssdc=1&sidt...Dr0l71t5NtKDNvgzI	302 Moved Temporarily	accounts.youtube.com	212 B	10.64.22.15:8080	112ms
GET ?auth=DQAAAIMAAAAf...Dr0l71t5NtKDNv	302 Moved Temporarily	mail.google.com	0	10.64.22.15:8080	152ms
GET ?shva=1	200 OK	mail.google.com	21.8 KB	10.64.22.15:8080	432ms
GET ?ui=2&view=js&name...k1HFMewXo6MJ(	200 OK	mail.google.com	343 KB		37ms
GET ?ui=2&view=bsp&ver=ohhl4rw8mbn4	200 OK	mail.google.com	62 B		35ms
GET ?ui=2&view=bsp&ver=ohhl4rw8mbn4	200 OK	mail.google.com	62 B		238ms
GET ?ui=2&view=bsp&ver=ohhl4rw8mbn4	200 OK	mail.google.com	62 B		239ms
GET ?ui=2&view=ss&mset...MewXo6MJQhTjE	200 OK	mail.google.com	55.2 KB		184ms
GET ?ui=2&view=jsm&nam...k1HFMewXo6MJ	200 OK	mail.google.com	5.4 KB		766ms
GET ?ui=2&view=ss&mset...MewXo6MJQhTjE	200 OK	mail.google.com	55.2 KB		688ms
GET sem_8e56e5be46cb600be9ba1b375de5d	200 OK	ssl.gstatic.com	12.1 KB		598ms
POST ?ui=2&ik=19011efaa...&rt=j&search=i	200 OK	mail.google.com	1.7 KB	10.64.22.15:8080	862ms
GET cleardot.gif?zx=1qmlhw87jnu	200 OK	mail-attachment.googleusercontent.com	43 B	10.64.22.15:8080	44ms
GET setgmail?zx=pc36swyh8rif	204 No Content	google.com	0	10.64.22.15:8080	70ms
GET ?s=gmail&a=viewinb...s&zx=1md35xr0y	200 OK	clients2.google.com	35 B	10.64.22.15:8080	475ms
GET ?ui=2&ik=19011efaa...k=W1UHOX3tnF9I	200 OK	mail.google.com	890 B	10.64.22.15:8080	159ms



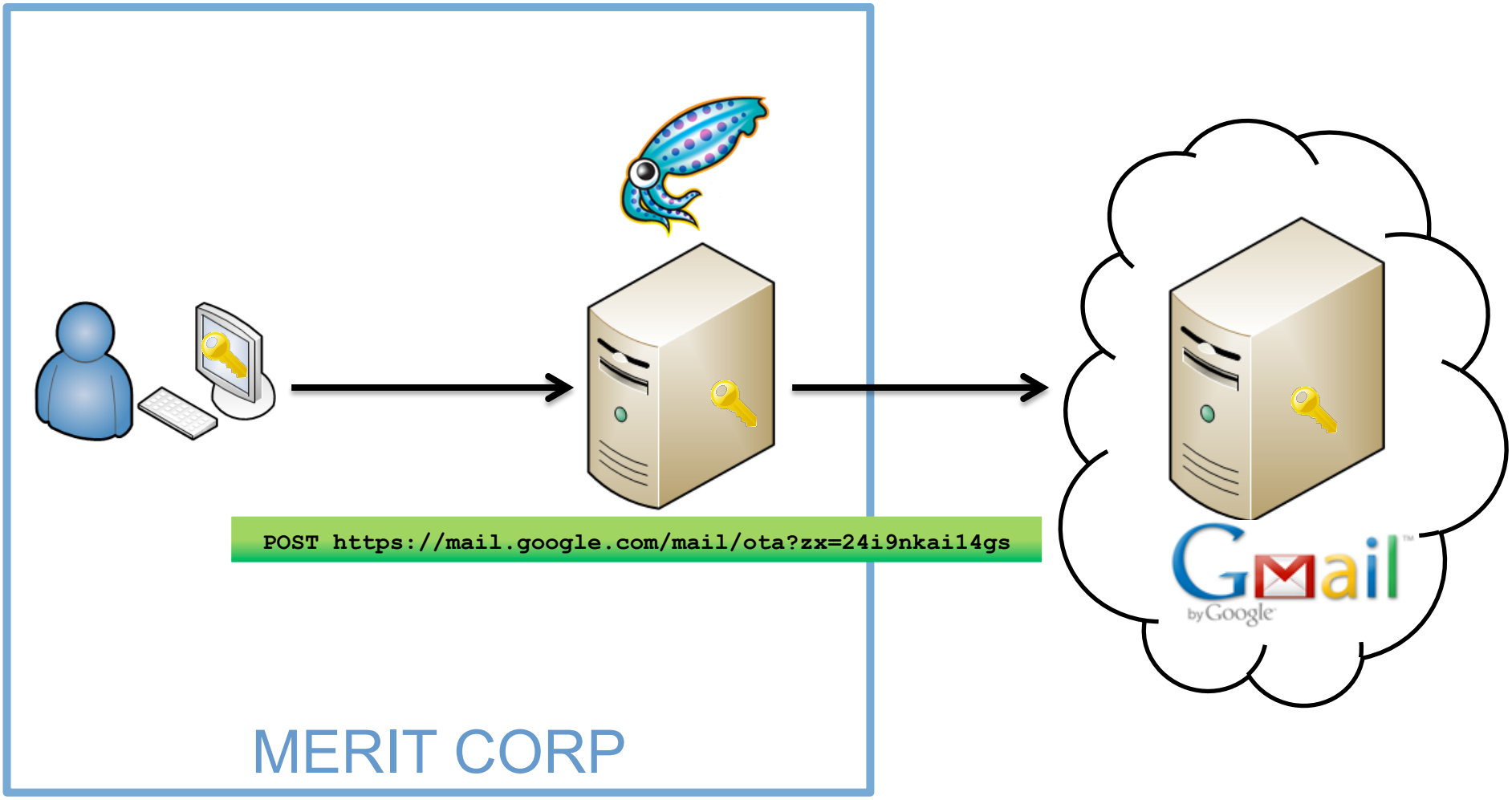
Man-in-the-Middle (MITM) Proxy



Squid's HTTP Request Logging

```
 image/gif
1331070430.915    101 10.0.3.100 TCP_MISS/200 491 GET https://mail.google.com/mail/images/c.gif? - DIRECT/74.125.225.86 image/gif
1331070432.096    160 10.0.3.100 TCP_MISS/200 502 POST https://mail.google.com/mail/ota? - DIRECT/74.125.225.86 text/plain
1331070432.894   2115 10.0.3.100 TCP_MISS/200 485 GET https://mail.google.com/mail/channel/test? - DIRECT/74.125.225.86 text/plain
1331070433.281    166 10.0.3.100 TCP_MISS/200 650 POST https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070433.948    226 10.0.3.100 TCP_MISS/200 930 GET https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070433.950    225 10.0.3.100 TCP_MISS/200 439 POST https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070433.958   1684 10.0.3.100 TCP_MISS/200 1488 POST https://mail.google.com/mail/? - DIRECT/74.125.225.86 text/javascript
1331070434.181    114 10.0.3.100 TCP_MISS/200 665 GET https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070434.224    204 10.0.3.100 TCP_MISS/200 816 POST https://mail.google.com/mail/? - DIRECT/74.125.225.86 text/javascript
1331070436.859    171 10.0.3.100 TCP_MISS/200 441 POST https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070438.694    166 10.0.3.100 TCP_MISS/200 501 POST https://mail.google.com/mail/ota? - DIRECT/74.125.225.86 text/plain
1331070438.811     12 10.0.3.100 NONE/403 905 POST https://mail.google.com/mail/? - NONE/- text/html
1331070440.557    174 10.0.3.100 TCP_MISS/200 441 POST https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070450.612  16313 10.0.3.100 TCP_MISS/200 638 GET https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070450.695   1816 10.0.3.100 TCP_MISS/200 1426 POST https://mail.google.com/mail/? - DIRECT/74.125.225.86 text/javascript
1331070477.220  26566 10.0.3.100 TCP_MISS/200 521 GET https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070485.588    180 10.0.3.100 TCP_MISS/200 441 POST https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070502.412  25158 10.0.3.100 TCP_MISS/200 521 GET https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070530.693  28245 10.0.3.100 TCP_MISS/200 521 GET https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070545.425    170 10.0.3.100 TCP_MISS/200 441 POST https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070557.055  26324 10.0.3.100 TCP_MISS/200 521 GET https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070583.442  26353 10.0.3.100 TCP_MISS/200 521 GET https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070605.601    328 10.0.3.100 TCP_MISS/200 441 POST https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.85 text/plain
1331070608.375  24891 10.0.3.100 TCP_MISS/200 521 GET https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070633.368  24967 10.0.3.100 TCP_MISS/200 521 GET https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070659.009  25609 10.0.3.100 TCP_MISS/200 521 GET https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070665.453    167 10.0.3.100 TCP_MISS/200 441 POST https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.85 text/plain
1331070685.258  26205 10.0.3.100 TCP_MISS/200 521 GET https://mail.google.com/mail/channel/bind? - DIRECT/74.125.225.86 text/plain
1331070686.760    145 10.0.3.100 TCP_MISS/200 906 POST http://safebrowsing.clients.google.com/safebrowsing/downloads? - DIRECT/74.125.225
```

Man-in-the-Middle (MITM) Proxy

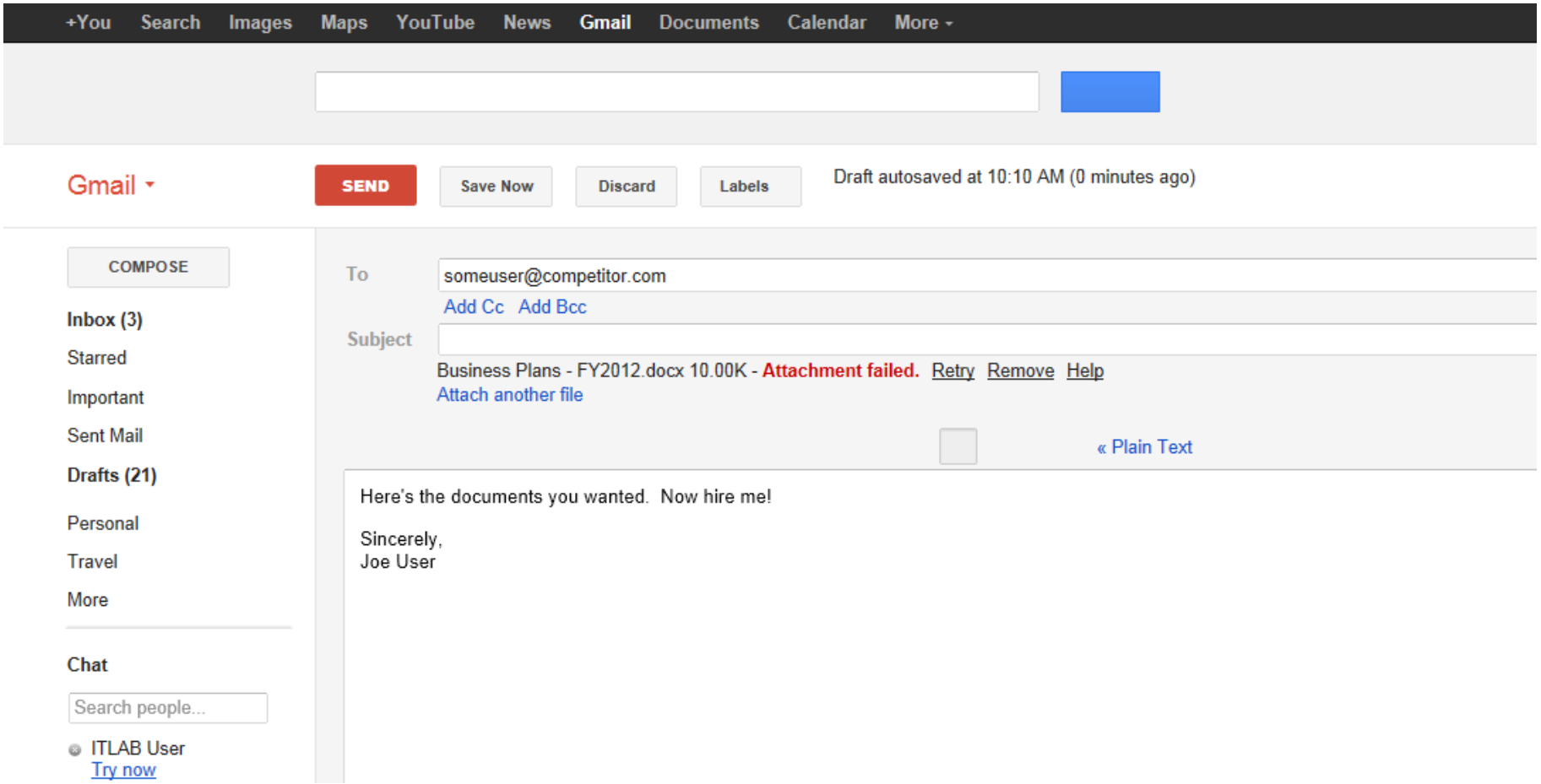


Man-in-the-Middle (MITM) Proxy

RegEx: `mail.google.com/mail/ota*`

```
POST https://mail.google.com/mail/ota?zx=24i9nkai14gs
```

Success!



Shortcomings

- Not very granular
- Doesn't account for the scenario where text is copied and pasted into an email

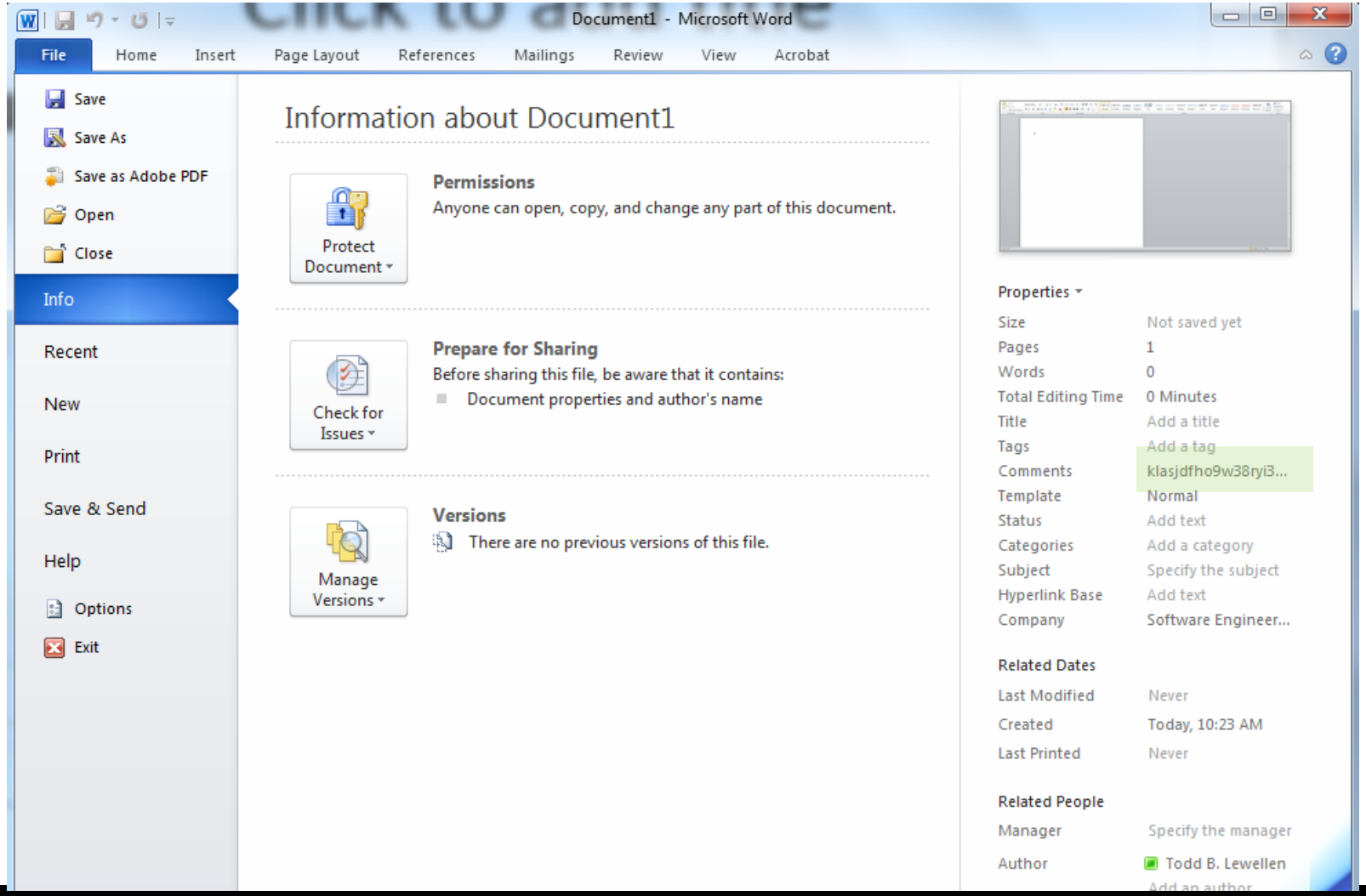
Detection using ClamAV

```
testSig:0:*:                    For Official Use Only
```

Detection using ClamAV

```
klasjdfho9w38ryi3ubsdkvjlaw3oy5423uihtgi
eaufsdlair78230895r82375g2389q7r834789hf
                kld3938fnf-
                          33437383968666b
```

Detection using ClamAV



Plagiarism Detection & DLP

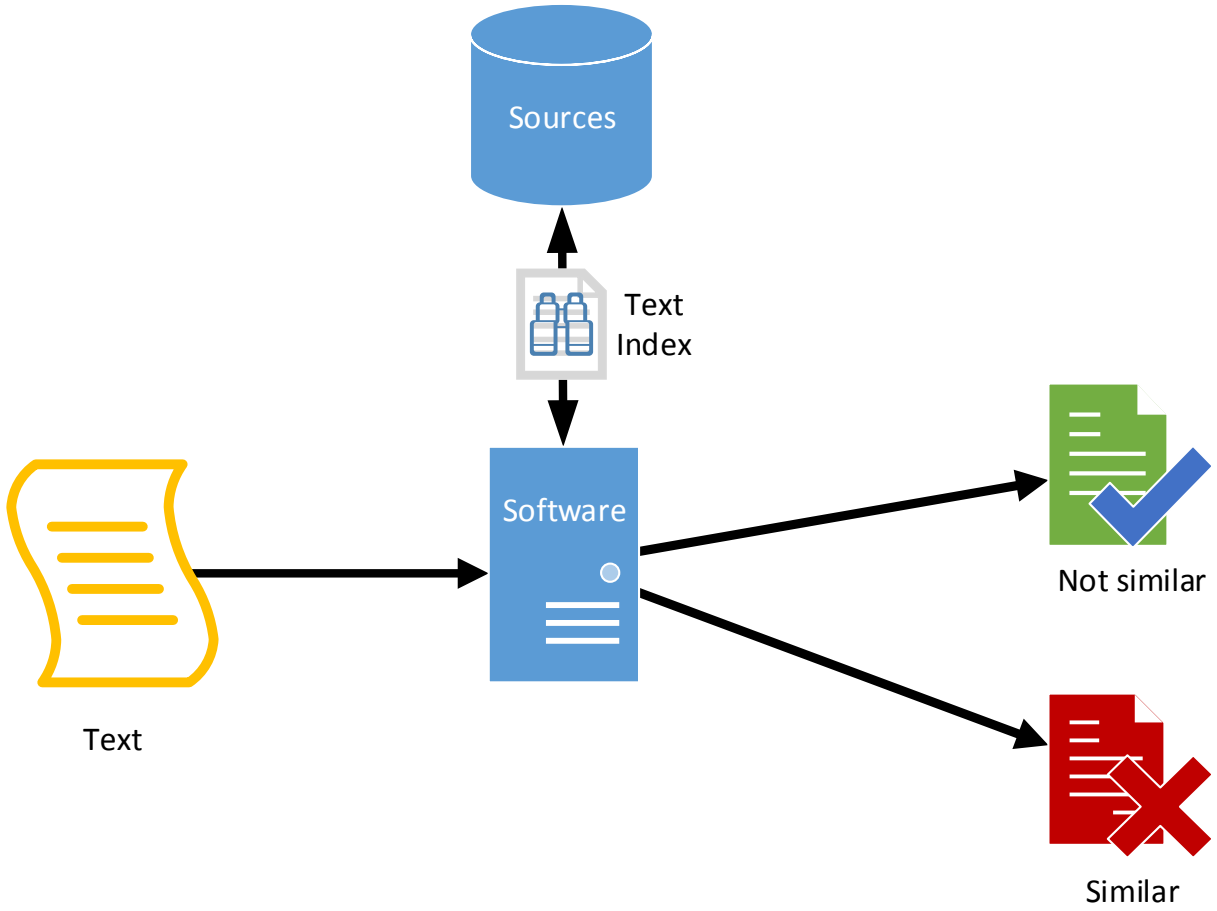
Solution:

- What if we could inspect all text flowing through the network?
- Rather than look for ‘tags’ or keywords, look for *similarity*
- How do we test document similarity?
- **Cosine similarity algorithms**
  - Laymen’s terms: Plagiarism Detection
  - Even though we’re not checking for plagiarism in academic papers, the process is virtually identical

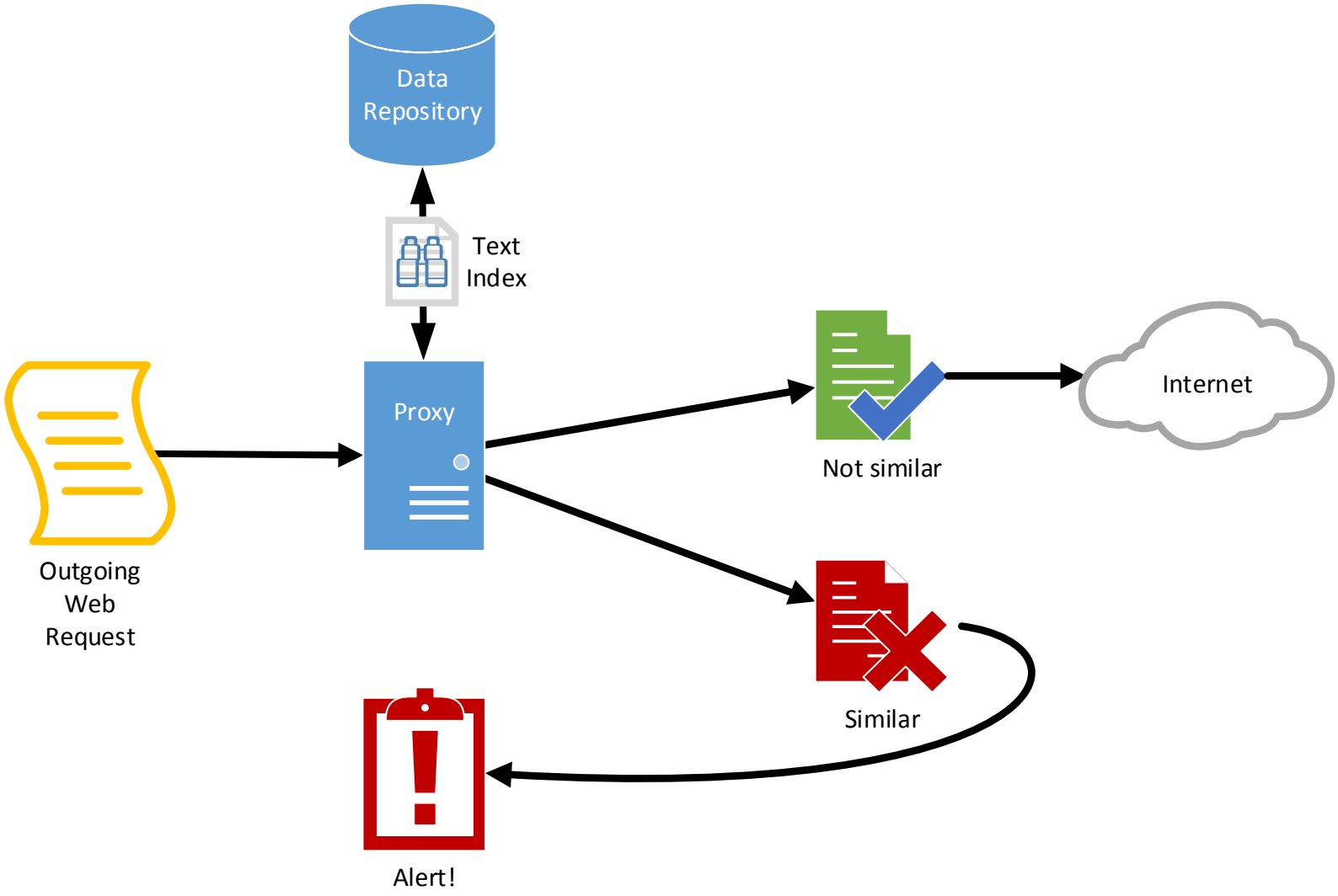
The Plagiarism Detection Method

- Rather than asking
  - "Does any text in this document sufficiently match anything within its cited references?"
- We're asking
  - "Does any text in this outgoing network traffic sufficiently match anything within our repository of intellectual property?"
  - If not – send it through
  - If so – create an alert *and/or* actively block the traffic from leaving the organization's perimeter

Plagiarism Detection



Plagiarism Detection in DLP



Open Source Tools

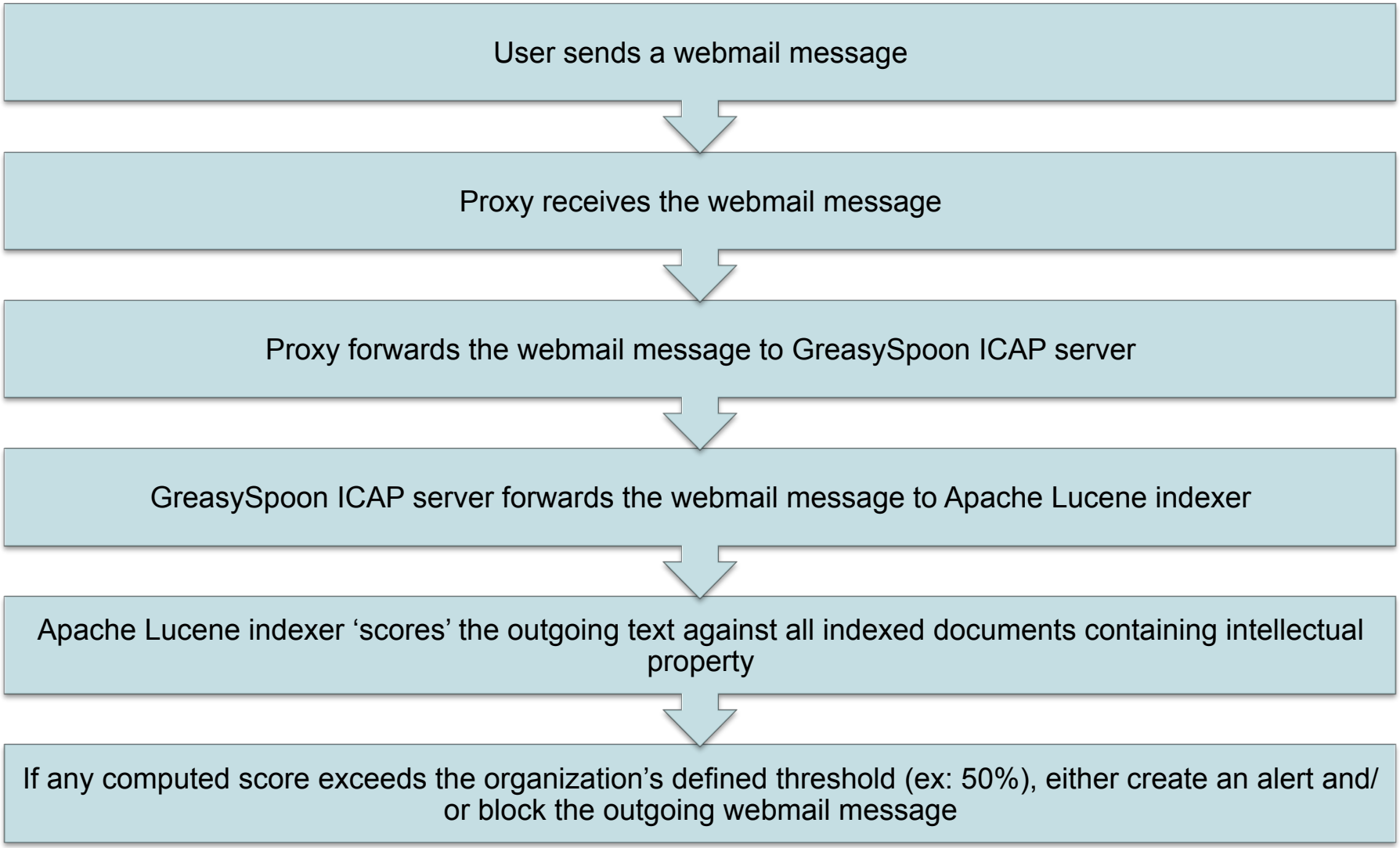
- Squid proxy server
- Apache Lucene
- Apache Tika
- GreasySpoon ICAP server

GreasySpoon

Apache Lucene

- Powerful open-source text indexer and search engine
- Used in IBM's famous Watson AI system
- Scalable, fast, and mature
- Perfect for our needs

Order of Events



Shortcomings

- Tuning the threshold is difficult
- Does not detect encodings other than ASCII or Unicode
- Processing intensive
- Large index (lots of duplicated data)
- Index contains sensitive information

Future Work

- Create an efficient open-source DLP framework for correlating any given input data with any set of data, regardless of their type (i.e. text, image, raw)
- Tagging network traffic with usernames and other attribution information
- Improving our “Tagger” tool to automatically store file usage information within documents when they are created/accessed/modified

Upcoming Control Topics

- Two Man Control For Operating Systems
  - Why is it so hard?
- Better Forensics for Insider Threat Indicators
  - How to use what we know more effectively

Point of Contact

Randy Trzeciak
Technical Manager, CERT Insider Threat

CERT Division
Software Engineering Institute
Carnegie Mellon University
4500 Fifth Avenue
Pittsburgh, PA 15213-3890
+1 412 268-7040 – Phone
rft@cert.org – Email

http://www.cert.org/insider_threat/

This material is based upon work funded and supported by the Department of Defense under Contract No. FA8721-05-C-0003 with Carnegie Mellon University for the operation of the Software Engineering Institute, a federally funded research and development center.

DM-0000556